The Advice Gap: Gender Disparities in Online Relationship Advice Communities

Henry Stanley
henry@henrystanley.com

Abstract

[150-250 words. Key points to cover:]

- Research question: Do men and women receive different advice in online communities?
- Method: 6,080 advice comments from 591 Ask Metafilter posts, LLM classification
- Main finding: Men receive 2.86x more critical advice (37.9% vs 17.6%, $\chi^2 = 282.14$, $p < 0.0001$)
- Robustness: Effect persists after controlling for severity, fault, problem type
- Validation: 96% agreement with human judgment on advice direction
- Implications: Help-seeking behavior, platform design

1 Introduction

[500 words. Key points to cover:]

- Online advice communities are widely used for personal problems
- Prior work on gender bias in online spaces (Wikipedia, Stack Overflow, harassment)
- Gap: Little research on whether *advice content* differs by recipient gender
- This study: Examine advice direction (supportive vs critical) and tone by poster gender
- Research questions:
  1. Do men/women receive different proportions of critical vs supportive advice?
  2. Do tone labels differ by gender?
  3. Do differences persist after controlling for confounds?

2 Related Work

2.1 Gender Bias in Online Communities

[200 words. Key citations:]

- Wikipedia gender gap [Lam et al., 2011, Hill and Shaw, 2013]
- Stack Overflow participation [Ford et al., 2016]
- Online harassment [Duggan, 2017, Pew Research Center, 2017]
- Feedback differences in professional contexts [Correll and Simard, 2016]

2.2 Advice-Giving Dynamics

[200 words. Key citations:]

- Supportive vs challenging advice [Goldsmith, 2004]
- Men and help-seeking, "tough love" norms [Addis and Mahalik, 2003]

2.3 LLM-Based Content Analysis

[200 words. Key citations:]

- LLMs for content analysis at scale [Ziems et al., 2024]
- Validation requirements [Gilardi et al., 2023]
- Performance on classification tasks [Törnberg, 2023]

3 Data and Methods

3.1 Data Collection

[200 words. Key facts:]

- Source: Ask Metafilter, "relationships" tag
- Known for active moderation, thoughtful responses
- Dataset: 591 posts, 7,091 comments (6,080 with advice)
- Gender distribution: 1,716 comments on male posts, 4,364 on female posts
- Imbalance reflects community composition

3.2 Classification Framework

[300 words. Describe classification scheme:]

Post-level variables:

- Poster gender (from explicit mentions, e.g., "I [30M]...")
- Situation severity: low / medium / high
- OP fault: none / some / substantial / unclear
- Problem category (communication, trust, boundaries, etc.)

Comment-level variables:

- Is advice: boolean
- Advice direction: supportive / critical / neutral / mixed
- Tone labels (12 total):
  - Positive: gentle, empathetic, constructive, understanding, encouraging, supportive
  - Negative: harsh, judgmental, blaming, dismissive, condescending, hostile

3.3 LLM Classification

[200 words. Key points:]

- Model: Claude Haiku 4.5
- Iteratively refined prompts
- Conservative criteria for negative tones (require clear evidence)
- Example: "judgmental" = explicitly condemning OP as bad person, not just pointing out mistakes

3.4 Validation

[150 words. Key points:]

- Human spot-checking of 51 random comments
- Advice direction: 96% agreement
- Tone labels: 57% agreement (more subjective)
- Re-classified with conservative criteria
- Primary analyses focus on high-reliability advice direction

3.5 Statistical Methods

[100 words. Methods used:]

- Proportions by gender
- Odds ratios with 95% CI
- Chi-square tests for independence
- Stratified analysis by severity, fault, category
- Two-tailed tests, $\alpha = 0.05$

4 Results

4.1 Dataset Characteristics

[100 words. Interpret the table below.]

[Key point: Severity and fault distributions do NOT differ by gender men and women post about comparable situations.]

4.2 Primary Finding: Advice Direction

[150 words. Interpret the table below.]

[Key statistics:]

- $\chi^2 = 282.14$, $p < 0.0001$
- Odds ratio for critical advice: **2.86** (95% CI: 2.50–3.27)
- Odds ratio for supportive advice: 0.41 (95% CI: 0.36–0.47)

Table 1: Post Characteristics by Poster Gender

Variable	Male (n=199)	Female (n=392)	χ^2	p
Severity			2.31	0.315
Low	18.1%	21.4%		
Medium	52.3%	48.7%		
High	29.6%	29.9%		
OP Fault			3.87	0.276
None	31.2%	35.7%		
Some	42.7%	38.5%		
Substantial	15.6%	17.1%		
Unclear	10.5%	8.7%		

Table 2: Advice Direction by Poster Gender

Advice Direction	Male	Female	Difference
Critical of OP	37.9%	17.6%	+20.3 pp
Supportive of OP	25.3%	45.3%	−20.0 pp
Neutral	24.1%	26.8%	−2.7 pp
Mixed	12.7%	10.3%	+2.4 pp

4.3 Tone Analysis

[200 words. Interpret the table below.]

[Key points: Men receive more negative tones (judgmental, blaming, harsh, condescending, hostile). Women receive more positive tones (understanding, empathetic, supportive, encouraging, gentle). Constructive and dismissive show no significant difference.]

4.4 Confound Analysis

[200 words. Interpret the stratified tables below.]

[Key points:]

- Effect persists across ALL severity levels (OR 2.4–2.9)
- Effect persists across ALL fault levels (OR 2.3–2.7)
- Notable: Men with NO fault (28.1%) receive more critical advice than women with SOME fault (20.4%)

4.5 Sensitivity Analysis

[100 words. Key points:]

- Core finding uses advice direction (96% human agreement) robust
- Excluding negative tone labels: positive tone differences remain significant
- Finding does not depend on subjective tone classifications

Table 3: Tone Labels by Poster Gender

Tone	Male	Female	Diff	χ^2	p
Positive tones					
Understanding	64.9%	72.9%	−8.0 pp	38.7	<0.0001
Empathetic	54.8%	65.5%	−10.7 pp	60.2	<0.0001
Constructive	47.3%	49.1%	−1.8 pp	1.7	0.19
Supportive	25.3%	45.3%	−20.0 pp	218.4	<0.0001
Encouraging	23.7%	34.5%	−10.7 pp	70.1	<0.0001
Gentle	12.4%	18.2%	−5.8 pp	32.1	<0.0001
Negative tones					
Judgmental	11.7%	4.0%	+7.7 pp	125.8	<0.0001
Blaming	9.2%	2.2%	+7.0 pp	145.3	<0.0001
Harsh	6.1%	3.0%	+3.1 pp	30.8	<0.0001
Condescending	4.5%	1.9%	+2.7 pp	31.6	<0.0001
Hostile	2.1%	0.3%	+1.8 pp	42.9	<0.0001
Dismissive	3.8%	3.2%	+0.6 pp	1.3	0.25

Table 4: Advice Direction by Gender, Stratified by Situation Severity

Severity	Gender	% Critical	OR	p
Low	Male	29.8%	2.41	<0.001
	Female	15.0%		
Medium	Male	38.4%	2.92	<0.0001
	Female	17.1%		
High	Male	43.2%	2.78	<0.0001
	Female	21.3%		

5 Discussion

5.1 Summary of Findings

[100 words. Summarize: Men 2.86x more critical advice, persists after controls, women get more supportive/empathetic responses.]

5.2 Interpretation

[200 words. Possible mechanisms:]

- Commenter stereotypes ("tough love" for men)
- Differential accountability standards
- Writing style differences (but confound analysis suggests this doesn't fully explain)
- Community composition (Ask Metafilter skews female in-group favoritism?)

5.3 Implications

[200 words. Discuss implications for:]

- Help-seekers (men should contextualize critical feedback)

Table 5: Advice Direction by Gender, Stratified by OP Fault

OP Fault	Gender	% Critical	OR	p
None	Male	28.1%	2.53	<0.001
	Female	13.1%		
Some	Male	41.2%	2.71	<0.0001
	Female	20.4%		
Substantial	Male	52.7%	2.34	<0.001
	Female	31.8%		

- Communities/platforms (consider bias in advice-giving norms)
- Research (LLM methodology for large-scale advice analysis)

5.4 Comparison to Related Work

[100 words. Compare to:]

- Prior work on men receiving less emotional support [Addis and Mahalik, 2003]
- Effect size (OR $\approx$ 3) larger than some professional setting biases
- Possible explanation: anonymous online contexts reduce social desirability

6 Limitations

[300 words. Address:]

- Single platform (Ask Metafilter has specific norms/demographics)
- LLM classification (validated but imperfect)
- Selection effects (can't observe who doesn't post)
- Correlation not causation (can't identify mechanism)
- Binary gender only (excludes non-binary, undisclosed)
- Temporal scope (community norms may change)

7 Conclusion

[200 words. Key points:]

- Evidence of substantial gender disparities in online relationship advice
- Men receive critical advice at 3x the rate of women
- Pattern persists across situation types and after controlling for severity/fault
- Implications for help-seeking behavior and platform design
- Future directions: cross-platform replication, mechanism investigation, interventions

References

Addis, M. E. and Mahalik, J. R. (2003). Men, masculinity, and the contexts of help seeking. *American Psychologist*, 58(1):5–14.

Anthropic (2024). Claude 3.5 Haiku model card.

Correll, S. J. and Simard, C. (2016). Research: Vague feedback is holding women back. *Harvard Business Review*.

Duggan, M. (2017). Online harassment 2017. Pew Research Center.

Ford, D., Smith, J., Guo, P. J., and Parnin, C. (2016). Paradise unplugged: Identifying barriers for female participation on Stack Overflow. In *Proceedings of the 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, pages 846–857.

Gilardi, F., Alizadeh, M., and Kubli, M. (2023). ChatGPT outperforms crowd-workers for text-annotation tasks. *arXiv preprint arXiv:2303.15056*.

Goldsmith, D. J. (2004). *Communicating Social Support*. Cambridge University Press.

Hill, B. M. and Shaw, A. (2013). The Wikipedia gender gap revisited: Characterizing survey response bias with propensity score estimation. *PloS One*, 8(6):e65782.

Lam, S. T. K., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., and Riedl, J. (2011). WP:Clubhouse? An exploration of Wikipedia's gender imbalance. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, pages 1–10.

Pew Research Center (2017). Online harassment 2017.

Törnberg, P. (2023). ChatGPT-4 outperforms experts and crowd workers in annotating political Twitter messages with zero-shot learning. *arXiv preprint arXiv:2304.06588*.

Ziems, C., Held, W., Shaber, O., Lu, J., Levy, M., Lahav, G., et al. (2024). Can large language models transform computational social science? *Computational Linguistics*, 50(1):237–291.